Notes on the Talk: Data System for Living With a Star

W. Dean Pesnell
Nomad Research, Inc.
pesnell@NomadResearch.com
Presented November 13, 2001

SUMMARY

During this presentation some thoughts and ideas about the Living With a Star Data System were presented. Examples of the Sun-Earth Connection were discussed, along with the sampling problem of high solar activity. The use of limb-sounding data from the thermosphere and ionosphere will require a larger archive than the time series oriented *in situ* data of earlier thermosphere/ionosphere missions. Once a regular series of data is available, assimilated data sets can be developed to maximize the utility of the observations.

INTRODUCTION

The Sun-Earth Connection can be most easily seen when the outer layers of the Earth's atmosphere vary in concert with solar activity. This variation has two components, solar extreme ultraviolet radiation (EUV, wavelengths from 1 to 120 nm) and magnetospheric inputs energized by the solar wind. The former variation can be illustrated with an empirical fit of the dominant component of the minimum exospheric temperature (without the latitude and diurnal dependencies) to the time history of F10.7, a proxy of solar activity (Figure 1),

$$T_{\min} = 379 + 3.24\,\mathrm{F10.7}_{81} + 1.3\,(\mathrm{F10.7} - \mathrm{F10.7}_{81})\ \mathrm{K}. \tag{1}$$

(F10.7 is described in the caption of Figure 1 and $\mathrm{F10.7}_{81}$ is the 81-day centered average of F10.7.)

This correlation fit was derived from satellite drag data (Jacchia, 1971; Champion, *et al.*, 1985). Similar fits derived from ground-based measurements would not resolve the latitude structure of the temperature. MSIS, a widely used empirical model of the thermosphere described in Hedin (1987), has a fit of the exospheric temperature derived from a number of satellites and ground stations. Such a fit is still limited by a lack of sampling at high solar and magnetospheric activity. The solar activity sampling is shown in Figure 2, where a histogram of the fraction of F10.7 values is shown. Relatively few points are available with $\mathrm{F10.7} > 250$.

When the magnetospheric (or auroral) inputs are considered, you are adding another poorly measured global quantity. Correlation fits are often derived for limited ranges of Kp or Ap, two proxy indicators of auroral input. Kp is an average of the magnetic disturbance at 12 globally distributed midlatitude stations. Ap is a daily average of a related index. Either Kp or Ap can be used as a proxy for auroral input but neither is a quantitative measure of that input. Correlation fits of the temperature change due to auroral input that use Kp or Ap give reasonable values for small to moderate values of those indices. Large values of Kp and Ap are sufficiently rare that the fits are probably inaccurate during times of large auroral input.

Another consideration of empirical models is the epoch of correlation. MSIS uses data from 5 radar stations, 8 satellites, and several sounding rockets to derive its correlations. Many of these measurements were made between 1960 and 1985. Continuing use of the MSIS model implicitly assumes that the minor constituents of the upper atmosphere have remained constant. That is surely not the case. Carbon dioxide and water vapor have steadily increased in the mesopause region. Increased CO_2 will reduce the temperature of the mesopause and cool the thermosphere and ionosphere. An increase in water vapor changes the chemistry in the lower thermosphere, with so far unknown effects.

DATA FROM THE IONOSPHERE AND THERMOSPHERE

Space-based data from the ionosphere and thermosphere (I/T) have traditionally been in situ instruments such as mass spectrometers, Langmuir probes, E-field booms, etc. These data are time series similar to the heliospheric and magnetospheric data described in earlier presentations. They produce quantitative data the is limited in range of solar activity, duty cycle of instruments, and length of mission. This last point is driven by the in situ requirement. By focusing on in situ measurements you must fly a satellite with a perigee less than 500 km. As the lifetime of a mission decreases rapidly with decreasing perigee altitude, you are sacrificing duration for sampling. In situ measurements are essential to measure the electric field drivers of the I/T region, for which we have no remote sensing capability, and to make precise composition measurements with high spatial resolution.

Remote sensing of the I/T will produce both image and limb-sounding data. The image data is similar to that of solar studies, spectroscopically resolved images designed to understand some part of an atmosphere. The data storage requirement of the image data may not be as large as the solar case as the anticipated cadence of images is not as rapid as for the sun. Limb-sounding data has not been considered in previous discussions so some details are needed.

Limb sounding is one method of generating altitude profiles of atmospheric quantities from the inversion of observed radiation. Radiances measured as a function of the tangent altitude (the minimum altitude of the optical ray) are used to constrain a model of the atmosphere. The model is varied until the radiances generated by the model agree (more or less) with the observations. This technique is well known in the troposphere where infrared measurements of water vapor and CO_2 are used to generate the temperature profile ($T[P]$). The TIMED mission will apply several remote-sensing techniques to the mesosphere and I/T.

Limb-sounding data requires a considerable amount of processing time to make it available to a user. In the jargon of EOS, Level 3 data are the inverted profiles of T, P, and possibly composition. Level 2 data are the radiances with pointing information and physical units. Although Level 2 is useful, most people will want to access the profiles. It is important to keep the lower levels of data because, in the unlikely event that an error is discovered in the processing software, all of the profiles must be re-generated. It is incumbent on the user of the data to be aware of the reprocessing status and to not use invalid data. TIMED encourages users to contact the PI to verify the status of the data before releasing new results.

Inversion of data often uses ancillary information that must be stored somewhere. SABER on TIMED requires the assimilative NCEP models of the troposphere and stratosphere to generate the lower boundary of their model (and initial guesses at the models). As SABER scans the limb with a mirror, it can use geometry to calculate the altitude of the lower boundary (in km), but to know the pressure of that location requires an interpolation in the NCEP models. A complete database would store any ancillary information needed to generate profiles until a final version is completed.

ASSIMILATIVE MODEL DATA

Although data assimilation has been mentioned frequently in the DSP, the I/T brings out the possibility of a new type of assimilated data. When many measurements from heterogeneous instruments are available to describe the I/T, assimilative model data can be used to serve that data in a form convenient to many users. Data does not validate this type of assimilative model, the model "interpolates" the data to a convenient spatial grid and set of time intervals. Assimilated model data gives global context to local data.

Model data of this form is used to serve tropospheric data to a wide audience, a good example is the NCEP data mentioned above. Data from many source, ground-launched balloons, commercial aircraft, and satellite-based sounders, are ingested into a specially designed model. The model is forced to agree with the data where information is available and uses a set of primitive equations (based on the equations of conservation of mass, momentum, and energy) to propagate the information to regions of the globe where measurements are lacking. At regular intervals the model output is stored at a data archive, which can then be accessed by user to initialize their models or for comparisons with forecasts. Individual users do not have to worry about the sources of the data, they can treat the assimilated data as ground truth.

If used in the LWS archive the amount of storage required begins to increase rapidly. Both the measured and assimilated data must be stored and served. As this is a goal of the LWS Program, it must be allowed for in the data archive. These assimilative models do not exist at present. Their development may be part of the Theory, Modeling, and Data Analysis component of LWS but their implementation would become part of the Data System.

DATA ARCHIVE CATALOG

I envision that the catalog used to search and organize the data is an essential part of the LWS data system. With a sufficiently advanced catalog, it is not necessary to start with a central archive—only the address of the data must be known. Much like the current way that libraries act as archives of government publications, universities could act as archive servers of data.

Research in searching heterogeneous datasets, such as the Internet, is ongoing among computer scientists. LWS should be able to use the results of this research to develop a capability to search and catalog a large set of observations from existing database software. Keeping new information about catalogued datasets, such as publications using that data and new correlations between previously unlinked data, should become part of the catalog. By working with a flexible database program, metadata may become heuristic, with the catalog learning new terms that can be applied to existing datasets.

Living With a Star should concentrate on serving the data in a flexible fashion, with the long-term archive only a part of the database. Mission data is a living entity, changing with revisions in algorithms and unsuited to a permanent archive until much work has been done with the data. As with the early data, correlation studies between the solar and magnetospheric drivers of the ionosphere and thermosphere will be an important part of Living With a Star. A flexible database will allow these studies to take place.

REFERENCES

Champion, K. S. W., A. E. Cole, and A. J. Kantor, Standard and reference atmospheres, in *Handbook of Geophysics and the Space Environment*, ed. A. S. Jursa, Springfield, VA, National Technical Information Service, 14-1–14-43, 1985.

Hedin, A. E., MSIS-86 thermospheric model, *J. Geophys. Res.*, **92**, 4649–4662, 1987.

Jacchia, L. G., Revised static models of the thermosphere and exosphere with empirical temperature profiles, *Smith. Astrophys. Obs. Spec. Rep.*, **332**, 1971.

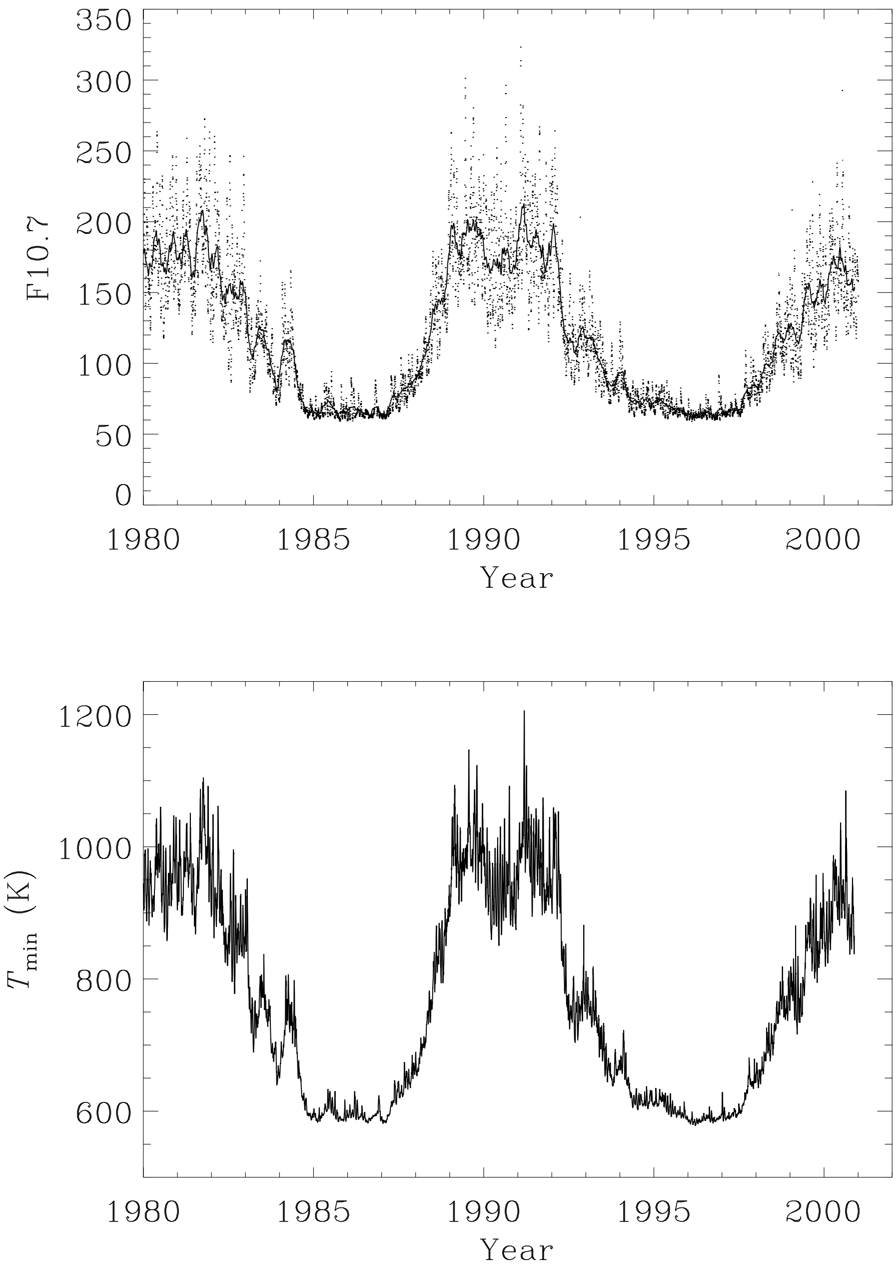

Figure 1: (Top) the solar irradiance at a wavelength of 10.7 cm (in units of 10^{-22} W m^{-2} Hz^{-1}) plotted as a function of time since 1980. The daily value is shown as a point and the 81-day centered average is shown as a solid line. This irradiance, called F10.7, is often used as a proxy for solar activity. (Bottom) An empirical fit for the minimum temperature in the exosphere (which is also the isothermal thermospheric temperature) derived from satellite drag data. The functional form of the fit is given in eq. 1.

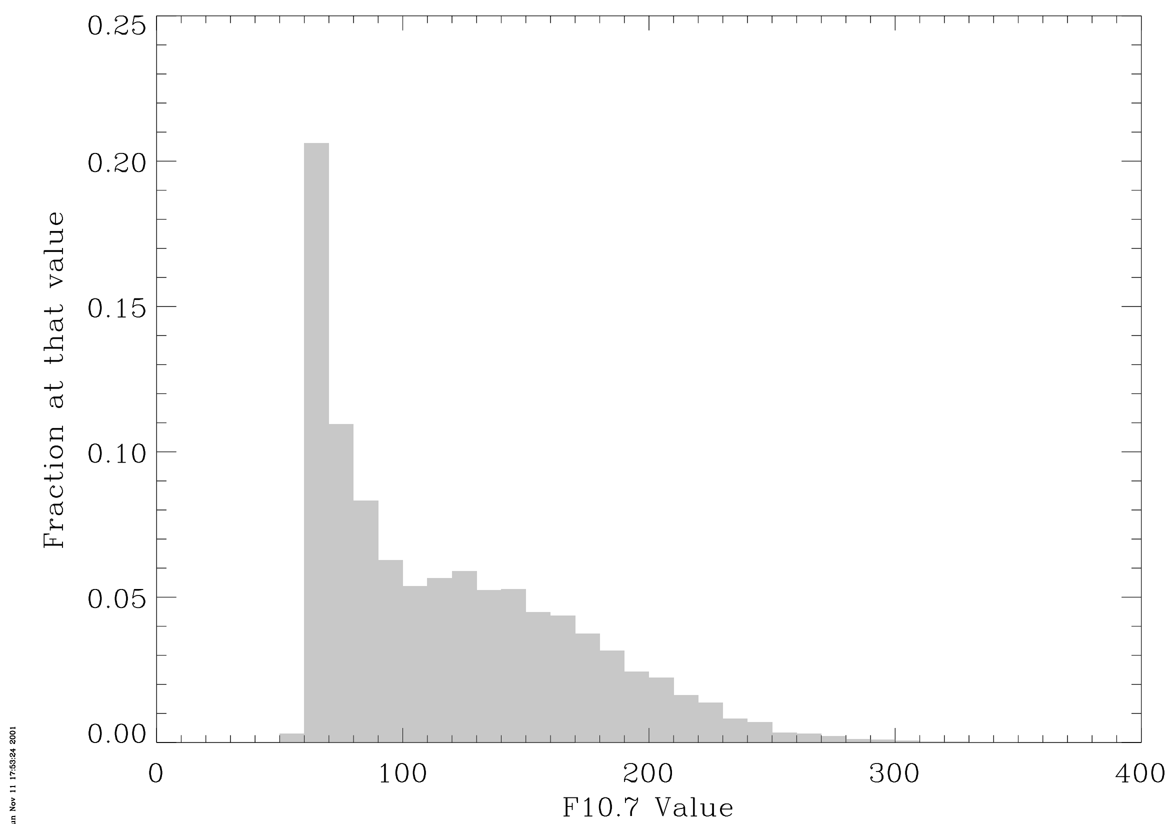

Figure 2: A histogram of F10.7 values from Figure 1 as a function of F10.7. The bins are 10 F10.7 units wide.